DOCUMENT RESUME

ED 143 684 TM 006 436

AUTHOR	Stein, Joan Lerner; And Others
TITLE	The Develcpment of Categorical Clustering: Validation of a New Measure. RIEP Print No. 68.
INSTITUTION	Research Inst. for Educational Problems, Cambridge, Mass.
SPONS AGENCY	National Inst. of Child Health and Human Development (NIH), Bethesda, Md.; Public Health Service (DHEW), Arlington, Va.
REPORT NO	RIEP-P-68
PUB DATE	May 75
GRANT	NICHHD-1-RO1-HDC8307
NOTE	17p.; For related document, see TM 006 441
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS	Age Differences; *Cluster Analysis; *Cluster Grouping; Cognitive Development; Data Analysis; Elementary Education; *Learning Processes; *Measurement Techniques; Mediation Theory; *Recall (Psychological); Sex Differences; Task Analysis

ABSTRACT

Research on 128 children is presented as evidence of the superiority of a new clustering measure (X) over a more traditional "chance clustering" measure (sigma) as a means of organizing material to be learned tc facilitate recall. X is shown to meet three criteria for a measure of a developmental process: X is more highly correlated with recall than sigma: X is more sensitive than sigma to the relationship between age and experience in the development of the clustering strategy: and X is sensitive to the partitioning of the stimulus list while sigma is not. X is discussed in terms of its ability to concurrently take into account both input and output dimensions of the task. By ignoring the qualities of the list to be learned, "chance clustering" measures treat clustering only as a phenomenon of the recalled list rather than as an adaptive strategy evaluated within the context of the dimensions of the task as it was presented. Research subjects were 128 elementary school students who performed recall tasks. (Author/MV)

****
* Documents acquired by ERIC include many irformal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED143684

PRINTS

THE DEVELOPMENT OF CATEGORICAL CLUSTERING: VALIDATION OF A NEW MEASURE

Joan Lerner Stein, Richard E. Crandall, and Milton Budoff

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Milton Budoff

RIEP-Print Number 68

May, 1975

RIEP-Prints are a continuation of the series STUDIES IN LEARNING POTENTIAL

RESEARCH INSTITUTE FOR EDUCATIONAL PROBLEMS, INC.

29 WARE STREET • CAMBRIDGE, MASSACHUSETTS 02138 • TELEPHONE (617) 868-0360-2

© 1975

TM006 436

Abstract

Research is presented validating the superiority of a new clustering measure (X) over a more traditional "chance clustering" measure (sigma). X is shown to meet three criteria for a measure of a developmental process: (a) X is more highly correlated with recall than sigma, (b) X is more sensitive than sigma to the relationship between age and experience in the development of the clustering strategy, and (c) X is sensitive to the partitioning of the stimulus list while sigma is not. X is discussed in terms of its ability to concurrently take into account both input and output dimensions of the task. By ignoring the qualities of the list to be learned, "chance clustering" measures treat clustering only as a phenomenon of the recalled list rather than as an adaptive strategy evaluated within the context of the dimensions of the task as it was presented.

The Development of Categorical Clustering: Validation of a New Measure

In recent years, a large literature relating to the development of behavior mediated by organized verbal representation has appeared. The ability to use verbal mediation is discussed by many including Piaget, Flavell, Jensen, Rohwer, Kendler, etc., as a critical qualitative shift at the base of complex, abstract cognitive abilities. A free recall task has been widely employed in the study of this development. Stimulus sequences in which items are conceptually related by membership in common classes are used in these studies. The "tendency" for items which are somehow related to one another to be recalled together has been labeled "clustering" (Shuell, 1969). It is hypothesized that clustering reflects the use of an organizational strategy made on the basis of categorical associations which function as mediators for efficient learning. Organizing material to be learned into such conceptual clusters facilitates recall.

The clustering paradigm is being used to ask critical questions about the nature of higher order, mediated learning. However, there have been obstacles to the systematic evaluation of this basic research. A universally acceptable measure of clustering derived from recall data has not yet been agreed upon (Cole, Frankel, & Sharp, 1971; Colle, 1972; Frender & Doubilet, 1974). While it is generally assumed that clustering increases with recall, no measure that treats clustering as

a cognitive strategy has been developed to date.

The purely mathematical roots of clustering theory are to be found in the statistical treatment of run moments by Mood (1940). But the measurement of clustering in free recall involves more variables than those treated in a runs analysis. Cohen, Sakoda, and Bousfield (Note 1) introduced the recall variable (total number of items recalled) into the calculation of clustering probabilities. Bousfield and Bousfield (1966) redefined this parameter by including the actual partitioning (number of categories and items) of the recall list in the calculation as well. Their measure, essentially the same as that used by Dallet (1964) is called "chance clustering." However, this measure also has shortcomings which hinder its usefulness to measure a cognitive process.

"Chance clustering" suffers from the defect that the actual stimulus category numbers, which are not the same as the recalled category numbers, do not enter into the calculation of clustering. "Chance clustering" does not take the organization of the stimulus list into account and therefore cannot be viewed as an indicator of a cognitive strategy. Secondly, there is the difficulty of extreme negative skewness of the data (Cole et al., 1971). As Shuell notes (1969), the range of such measures cannot be symmetrical about zero and also independent of the maximum number of clusters. It is nearly impossible to compare clustering ability for different

recall protocols using such linear measures because researchers in the past have failed to use linear transformation of the data. The difficulty is the model-independent nature of the assumption that the ability to cluster is, a posteriori, the ability to recall.

Since recall and clustering are so correlated in a functional sense, the authors have developed a new measure to study clustering which takes this into account. The new parameter, X, is concerned not only with the categories recalled but takes into account the category structure of the stimulus list as well. The partitioning of the stimulus list is viewed as critical when studying clustering as a cognitive strategy since the concern is with the ability to reorganize material that is critical rather than just the characteristics of what is recalled. As noted, older measures (exemplified by the "chance clustering" measure, sigma (Hudson & Dunn, 1969)) ignore the presentation partition and depend only on clusters within the recall lists.

This paper reports data which compare the utility of X, a measure which considers the structure of the presentation and recall lists, and sigma, a measure of chance clustering which considers only the structure of the person's recalled list. We hypothesized that:

1. X scores shall vary as a function of the structure of the presentation list, unlike the measure of chance clustering.

2. While both clustering measures should increase

developmentally with chronological age and experience on the task, X should be a more sensitive indicator of such changes.

3. X should be more sensitively related to recall because it considers the proportion of the stimulus list recalled rather than deriving the clustering score from the items recalled without regard to the task as presented.

Method

Subjects

The subjects were 128 white middle class children from a suburban Boston community. Sixty-four boys and 64 girls participated in the study with each sex equally represented in the second, third, fourth, and fifth grades (32 subjects per grade). All subjects were of average intelligence. The mean IQs (SDs) of subjects in each grade were as follows: Grade 2, 116.74 (12.71), Grade 3, 115.20 (11.33), Grade 4, 119.45 (10.81), Grade 5, 112.93 (14.40).

Materials

Cardboard-mounted black and white pictures of familiar objects served as the stimulus materials. Two eighteen-item lists were constructed. List A consisted of three categories with six items in each category, and List B consisted of six categories with three items per category. The categories represented in Lists A and B were selected from those used in other studies and included clothing, musical instruments, animals, parts of the human body, fruit, furniture, tools, eating utensils, and vehicles. Six random order presentations of each list were constructed.

Procedure

Each subject was seen individually for about ½ hour. A study-test procedure was used for six trials. One half of the subjects, equally distributed by grade and sex, were randomly selected to be tested on List A and one half on List B. Subjects were shown each of 18 stimulus pictures for five seconds during the study phase and were then asked to recall as many pictures as they could during the test phase of each trial. A different random order of the list was presented for each of the six trials.

Data Analysis

Four-way repeated-measures analyses of variance were performed on three dependent measures: recall, X, and sigma, a measure of chance clustering. The formula for recall was $R = r/N$, where r equals the number of items recalled in a trial and N equals the number of words in the stimulus list. A recall score of 1 represents a perfect score. In this study N equals 18.

X was calculated using the following formula: $X = (CL/r) - 1$. In this formula, r equals the number of items recalled (Rn), L equals the number of categories in the stimulus list (three for List A and six for List B), and C equals the sum of a series of constants multiplied by the number of clusters of different sizes. X has unit variance. For all possible partitions, X cannot be greater than 13.0 or less than -3.3. A full discussion of the derivation and other properties of X can be found in Crandall (Note 2).[1]

The formula used to calculate sigma for each trial was:

$$\frac{r + 1 - (1/r) * 364/L^1 - B}{r + 1 - (1/r) * 364/L^1 - L^1} = \sigma \qquad (1)$$

where r equals the number of items recalled (Rn), L equals the number of recalled categories, and B equals the number of clusters in the recalled list. Sigma does not have unit variance. The range for scores is -1 to 1. Crandall presents an analysis of this formula in his 1973 paper (Note 2).

The four factors used in the analyses were grade (2, 3, 4, and 5), sex (male and female), list (List A with three categories and List B with six categories), and trials (1 through 6). The first three factors were between-subjects factors while trials was a within-subjects repeated measure.

Results

In this study it was hypothesized that X would be a superior measure of clustering as a learning strategy due to its sensitivity to the categorical structure of the material to be learned. When X was used as the dependent measure a significant List effect was obtained (F (1,112) = 6.217, p =.015). For all subjects, across-trials, more clustering was evident on List B (six categories, three items) than on List A (three categories, six items). The mean X for List A was 1.5 and the mean for List B was 1.9. By contrast, the analyses of the chance clustering measure, sigma, did not yield any differences between the two lists. The two equivalent means on sigma were .43 for both lists. The use of a

measure which considers list characteristics clearly provides more sensitive measurement of the occurrence of clustering.

A criterion for the value of any clustering measure would be a confirmation of the developmental expectation that clustering would increase with age and with experience. It was hypothesized that $\underline{X}$ would be more sensitive to the developmental shifts in the use of clustering than sigma. Significant main effects for Grade ($\underline{F}$ (3, 112) = 7.880, $\underline{p}$ <.001) and Trials ($\underline{F}$ (5, 560) = 32.887, $\underline{p}$ <.001) indicated that clustering as measured by $\underline{X}$ improved in the expected directions. As might be expected, both Recall and Sigma also increased significantly across Age and Trials. However, a significant Grade X Trials interaction was also obtained ($\underline{F}$ (15, 560) = 2.863, $\underline{p}$ <.001) for $\underline{X}$. It can be seen from Figure 1 that clustering appears to be unrelated to Grade at Trial 1. On Trial 6 there is a gradual increase in mean score from grades 2 to 5 with the largest discrepancy between grades 3 and 4. Older children are more apt to improve their clustering scores across trials than younger children and to achieve higher scores as well. When sigma is used to measure clustering, this developmental shift is not found.

Insert Figure 1 about here

Although the Grade X List interaction on $\underline{X}$ was not significant, the means indicate that the higher $\underline{X}$ score for List B occurs only in Grades 4 and 5. It is the older subjects,

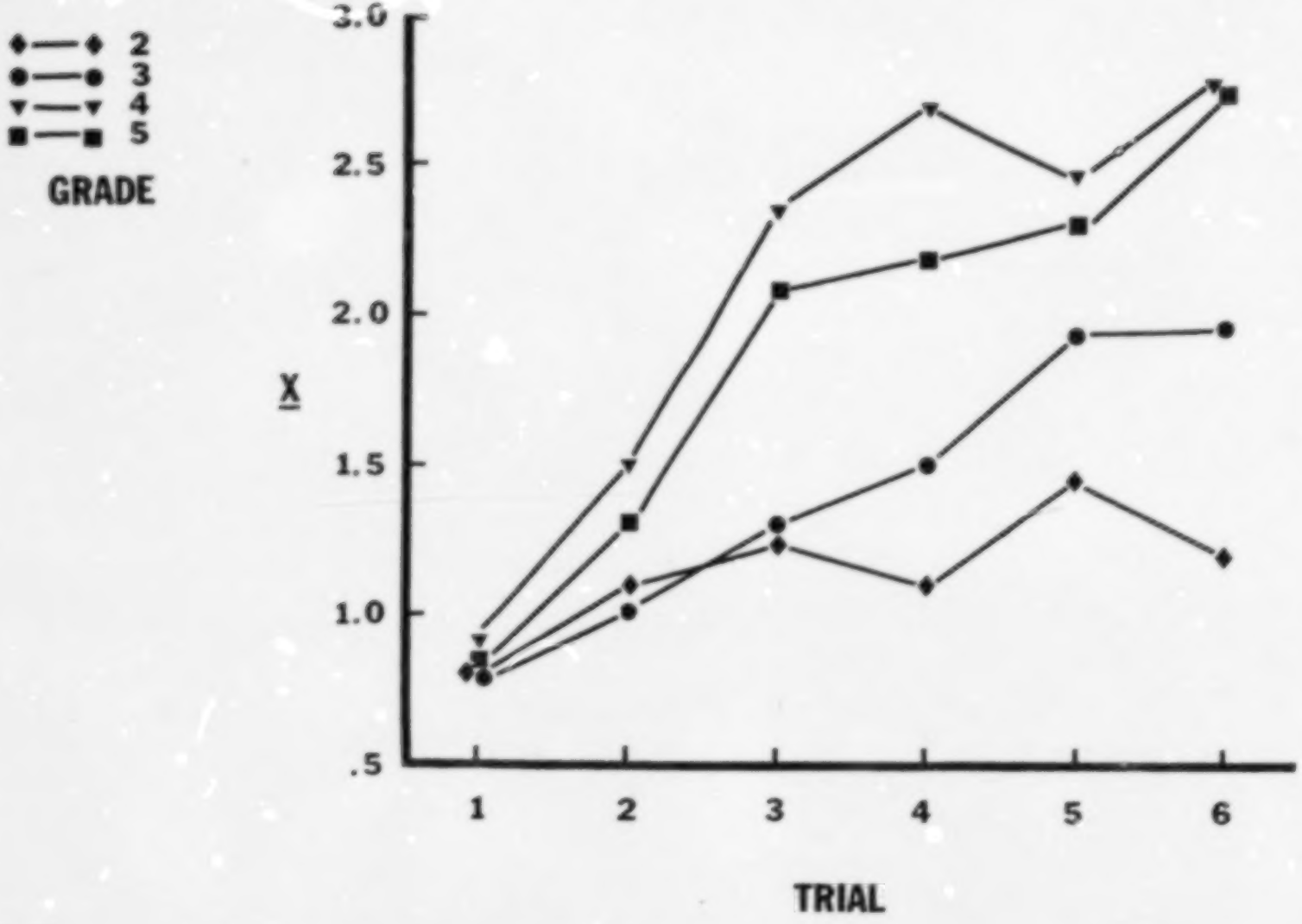

Figure 1. Clustering scores as measured by X on six trials for grades 2, 3, 4, and 5.

who do more clustering, who contribute to the differential X on the two lists. In the younger grades, when the amount of clustering is low, sensitivity to the list does not occur. Again, sigma scores do not follow this developmentally important pattern. Interactions among grade, list, and trial were not significant for either X or sigma. No sex effects or interactions were found for any of the dependent measures.

Finally, as hypothesized, X was significantly more highly correlated with recall than sigma. The measure X, taking into account the structure of the recall list relative to the presentation list, more directly measures clustering as a strategy to improve learning. For X and recall, the range of correlations was from .24 at Trial 1 to .6 at Trial 6. The correlation between sigma and recall was .14 at Trial 1 and .29 at Trial 6. The scatter diagram confirms the prediction from the derivation of the measures that sigma can be high when amount recalled is low while X cannot. The two measures are more equivalent in the middle range of recall scores but X is more sensitive to the extremes.

Discussion

Results of this study confirm the greater sensitivity of the measure X as an index of categorical clustering. Three criteria for a measure of a developmental process are met: (a) X is more highly correlated with recall than chance clustering, (b) X is more sensitive than sigma to the relationship between age and the ability to profit from experience, and (c) X is sensitive to the partitioning

of the stimulus list. X may be viewed as a measure which reflects the child's ability to take account of the organizational properties of the material to be learned as an important constituent element in applying a clustering strategy. We might expect that children of different ages will differ with regard to this ability and that younger children will be less likely to discover the organizational (categorical) cues embedded in the lists. This study using the measure X does indicate that the saliency of the categorical structure of a list to be learned is a critical factor in the emergence of the clustering strategy.

At about age 8, children appear to be able to profit from experience and discover the categorical nature of the lists. This discovery is then used to mediate the learning task. This finding, taken with the observation that older children are more apt to show superior clustering on the list with more salient categories while performance of younger children does not differentiate the two lists, lends further support for the significance of the organization of the list in the development of the strategy. While both X and sigma increase with age, X has the additional power to discriminate this critical characteristic of the list to be learned. Use of X as the clustering measure allows for more sensitive study of the psychological significance of partitioning in the development of the clustering strategy.

We have noted that traditional clustering measures deal

cnly with the characteristics of the recalled lists. A score is derived on the basis of the amount of items recalled either alone or in clustered combinations. The list from which items are to be recalled is not considered either in terms of length or in terms of its partition. Both length and partitioning impose limits on the possibilities for clustering. By ignoring the qualities of the list being recalled, chance clustering measures treat clustering only as a phenomenon of the recalled list rather than an adaptive strategy evaluated within the context of the dimensions of the task as it was presented. For this reason a perfect chance clustering score can be achieved if only a small part of the list is recalled as long as all items recalled are from one cluster. X on the other hand, is a measure which concurrently considers the input dimensions of the task and the child's responses. The higher correlation with recall further attests to the usefulness of this measure as an index of this cognitive strategy.

In this study we have tested the sensitivity of a new clustering measure to illustrate the development of the verbal mediational skill of clustering against more traditional measures of the use of this strategy. The clustering paradigm is a useful metaphor for studying cognitive development in a more general way. It is our belief that the parameter X will enable a more productive inquiry into these issues.

Reference Notes

1. Cohen, B. H., Sakoda, J. M., & Bousfield, W. A. The statistical analysis of the incidence of clustering in the recall of randomly arranged associates. (ONR Technical Report No. 10, Contract Nonr-631 (00)) Washington, D. C.: U. S. Government Printing Office, 1954.
2. Crandall, R. E. A theorem on clustering measures. RIEP-print No. 86, 1973. Research Institute for Educational Problems, 29 Ware Street, Cambridge MA 02138.

References

Bousfield, A. K., & Bousfield, W. A. Measurement of clustering and of sequential constancies in repeated free recall. Psychological Reports, 1966, 19, 935-942.

Cole, M., Frankel, F., & Sharp, D. Development of free recall learning in children. Developmental Psychology, 1971, 4, 109-123.

Colle, H. A. The reification of clustering. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 624-633.

Dallet, K. M. Number of categories and category information in free recall. Journal of Experimental Psychology, 1964, 66, 1-12.

Frender, R., & Doubilet, P. More on measures of category clustering in free recall--although probably not the last word. Psychological Bulletin, 1974, 81, 64-66.

Hudson, R. L., & Dunn, J. E. A major modification of the Bousfield (1966) measure of category clustering. Behavior Research Methods and Instrumentalism, 1969, 1, 110-111.

Mood, A. M. The distribution theory of runs. Annals of Mathematical Statistics, 1940, 11, 367-392.

Shuell, T. J. Clustering and organization in free recall. Psychological Bulletin, 1969, 72, 353-374.

Footnotes

This research was supported by grant 1 RO1 HD08307 from the National Institute for Child Health and Human Development, DHEW Public Health Service. The Research Institute for Educational Problems is located at 29 Ware Street, Cambridge MA 02138.

[1]An explanation of the derivation of $\underline{X}$ and a manual including tables to facilitate computation of $\underline{X}$ may be obtained from the authors at Research Institute for Educational Problems, 29 Ware Street, Cambridge MA 02138.